Characterization of a novel *POLD1* missense founder mutation in a Spanish population

Short running title: *Novel POLD1 missense founder mutation*

Rosario Ferrer-Avargues (1,2)*, Virginia Díez-Obrero (1,2)*. Ester Martín-Tomás (3). Eva Hernández-Illán (4,2), María-Isabel Castillejo (1,2), Alan Codoñer-Alejos (1,2), Víctor-Manuel Barberá (1,2), Ana-Beatriz Sánchez-Heras (5), Ángel Segura (6), María-José Juan (7), Isabel Tena (8), Adela Castillejo (1,2)**, José-Luis Soto (1,2)**.

1. Unidad Genética Molecular. Hospital Universitario Elche. Elche, Spain.
2. Instituto Investigación Sanitaria y Biomédica Alicante (ISABIAL). Alicante, Spain.
3. Análisis Clínicos. Hospital Universitario Elche. Elche, Spain.
4. Unidad Investigación. Hospital Universitario Alicante. Alicante, Spain.
5. Consejo Genético en Cáncer. Hospital Universitario Elche. Elche, Spain.
6. Consejo Genético en Cáncer. Hospital Universitario La Fe. Valencia, Spain.
7. Consejo Genético en Cáncer. Fundación Instituto Valenciano Oncología. Valencia, Spain.
8. Consejo Genético en Cáncer. Hospital Provincial Castellón. Castellón, Spain.

*Rosario Ferrer-Avargues and Virginia Díez-Obrero contributed equally to this work and share first authorship.

**Adela Castillejo and José-Luis Soto have contributed equally to this work and share senior authorship.

Corresponding author:

Name: José Luis Soto
Mailing address: Molecular Genetics Unit
Elche University Hospital
Camino Almazara nº 11
03203 Elche. Spain
e-mail address: soto_jos@gva.es
Telephone: 034 966616185
Fax: 034 966616127

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/jgm.2951

ABSTRACT

Background

We identified a new and a recurrent *POLD1* mutation associated with predisposition to colorectal cancer (CRC). We characterized the molecular and clinical nature of the potential *POLD1* founder mutation in families from Valencia (Spain).

Methods

Clinical and molecular data were collected from 4 independent families known to have a *POLD1* Leu474Pro mutation. To establish its founder effect, haplotype construction was performed using 14 flanking *POLD1* polymorphic markers. We calculated penetrance estimates and clinical expressivity, globally and stratified by age and sex.

Results

We included 32 individuals from the 4 families: 20 carriers and 12 noncarriers. A common haplotype was identified in these families in a region comprising 2.995 Mb, confirming L474P as the first founder *POLD1* mutation identified. Thirteen tumors diagnosed in 10 *POLD1* carriers: 8 CRC, 3 endometrial, and two other tumors were considered. The median age of cancer onset for *POLD1* mutation carriers was 48 years. The observed penetrance was 50% and the cumulative risk at age of 50 was 30%.

Conclusions

Our findings contribute to a better understanding of CRC genetics in the Spanish population. **The clinical phenotype for this mutation is similar to that in Lynch syndrome. Future studies using NGS with large gene panels for any hereditary cancer condition will offer the possibility to detect POLE/POLD1 mutations in unsuspected clinical situations showing a more real and unbiased picture of the associated phenotype.**

Keywords:

Cancer – colon/rectal

Molecular – genetics

INTRODUCTION

Polymerase proofreading-associated polyposis (PPAP) has been proposed as a new hereditary cancer condition with high-penetrance, dominant inheritance, characterized by multiple colorectal adenomas and carcinomas and caused by germline mutations in the genes encoding the catalytic subunits of the DNA polymerases Pol δ and Pol ε (*POLD1* and *POLE*, respectively)[1]. A major role of Pol δ in both the leading and lagging DNA strands has been described.[2]

This new entity emerged from the findings of The Cancer Genome Atlas project on colorectal cancer (CRC), in which a subset of ultramutated tumors with absence of microsatellite instability (MSI) and recurrent somatic mutations in *POLE* exonuclease domain was identified.[3] The number of germline *POLD1/POLE* proofreading domain mutations analyzed to date is relatively modest, and the majority of cases have been found to have microsatellite stable tumors,[1,4] even though MSI has been noted in some cases.[5] The most common germline mutations found to date are *POLE* L424V and *POLD1* S478N, although an increasing number of other different pathogenic variants are being identified.[1] The PPAP phenotype overlaps that of Lynch syndrome (LS) and *MUTYH*-associated polyposis (MAP), and the proposed screening and management algorithms are broadly similar. The analysis of *POLE* and *POLD1* is included in the diagnostic algorithm for these 2 cancer syndromes.[6]

We previously reported a new pathogenic *POLD1* variant (L474P) in 2 independent families classified as having familial CRC type X.[6,7] The finding of 2 other, apparently unrelated families from the same population drove us to establish its founder effect and its phenotype description.

MATERIALS AND METHODS

Patients, DNA, tissue samples, and clinicopathological information

Members from 4 nonpolyposis hereditary CRC families known to carry the *POLD1*-Leu474Pro mutation were recruited through the Hereditary Cancer Program of the Valencian region between 2005 and 2016. In addition, blood DNA from 30 healthy individuals was used to analyze the haplotype frequency in controls. Biological

samples and clinicopathological information were obtained from the Valencian Biobank Network and from the Hereditary Cancer Program of the Valencian region, both in Spain. The study was approved by the Ethics Committee of the Department of Public Health of the Valencian region.

Diagnostic algorithm of the nonpolyposis hereditary CRC

Clinical and pathological criteria (Amsterdam and Bethesda criteria) were used as the first tier for hereditary CRC suspicion. Alternatively, universal screening of CRC and endometrial cancer, diagnosed below the age of 71 years by immunohistochemistry (IHC) of the mismatch repair (MMR) proteins, is also considered. MMR status in tumor tissue was assessed either by IHC of the MMR proteins or by PCR-based MSI analysis, or both. *MLH1* promoter methylation and *BRAF*_V600E mutation in tumor tissues was analyzed when MLH1 loss of expression was detected. Families fulfilling Amsterdam II criteria with normal expression of MMR proteins or microsatellite stable tumors were considered as familial CRC type X, and genetic analysis of *POLE* and *POLD1* was performed. When loss of MMR protein expression in tumors was found, the germline analysis of the involved gene was assessed.

Germline mutation analysis

Germline mutation studies were performed on DNA isolated from peripheral blood leucocytes. Sanger sequencing was used to screen for mutations in exons where the most frequent mutations are located (exon 13 of *POLE*, and exon 11 of *POLD1*). Primers and PCR conditions are published elsewere.[7] Detection of point mutations in *MLH1* was conducted using PCR and direct sequencing of the whole coding sequence and intron–exon boundaries for each gene and large rearrangements analysis by MLPA.[8] Sequencing was performed using a 3130 Genetic Analyzer (Applied Biosystems, Foster City, CA, USA), and data were analyzed using Sequencing Analysis, version 5.1, and Variant Reporter, version 1.1, software (Applied Biosystems).

Microsatellite genotyping and haplotype analysis

Haplotypes were constructed manually from 3 SNPs (rs7246646, rs1363756, and rs12975011) and 11 *POLD1* flanking microsatellite markers (*D19S420, D19S902, D19S879, D19S867, D19S904, D19S907, D19S206, D19S921, D19S927, D19S891, D19S887*) expanding 13.818 Mb and assuming the lowest number of

recombinations. References, primer sequences, and PCR conditions are described in Table S1. SNP analysis was performed by allelic discrimination using real-time PCR with TaqMan probes (Applied Biosystems). Microsatellite marker analysis was performed with one fluorescently labeled primer per set according to standard procedures. Fragment analysis was performed using an ABI3130 analyzer with GeneMapper software (Applied Biosystems). Mutation age was estimated with the single marker method,[9] using affected and nonaffected chromosomes from the available members of the studied families.

RESULTS

We included 32 individuals from the 4 known families with the *POLD1*-Leu474Pro mutation: 28 were genetically tested individuals: 16 carriers and 12 noncarriers. In addition, 3 affected and obligated carriers were also considered. We exhausted the DNA from one mutated case that was excluded from the haplotype analysis.

Haplotype analysis and estimation of mutation age

We found a minimum common haplotype for all 16 mutation carrier individuals analyzed from the 4 families. The common haplotype extended from *D19S904 to D19S921* markers and represented a size of 2.995 Mb (Table 1, Table S2).

Thirty control individuals were genotyped for the microsatellite markers that define the common haplotype. We calculated the allelic frequency for each allele composing the minimum common haplotype. The estimated frequency of the common haplotype in controls was $1.25 \cdot \times 10^{-6}$ (Table S3). The resulting ages are 4 and 28 generations (100 to 700 years) for recombinant markers *D19S867* and *D19S927*, respectively.

Characterization of the clinical phenotype associated

The first family we found with this new mutation in *POLD1* was published by Valle et al.[7] This family meets Amsterdam II criteria, without alteration of the MMR system, being considered as having familial CRC type X. The mutation in this family was identified in 2 women affected by multiple tumors and in a healthy woman (Figure S1). The pathogenicity assessment of this mutation was established based on various lines of evidence. The Leu474 of *POLD1* is the paralog residue to the

Leu424 in *POLE*, which is clearly pathogenic. Moreover, it is highly evolutionally conserved, various in silico prediction programs suggest its pathogenicity, and functional in vitro studies (in yeast) show a loss of polymerase functionality, producing a mutator phenotype. The mutation cosegregated with the disease in the family.[7]

The second family found with this mutation was described by Bellido et al.[6] This family, apparently unrelated to the previous family, met Amsterdam I criteria, with both microsatellite stable (MSS) and MSI cases (loss of MSH6), in which the mutation has been identified in 7 individuals: 2 healthy and 5 diagnosed with CRC, and colonic and gastric polyps (Figure S2).

More recently, we have identified this mutation in a third independent family. This family consists of a woman who developed CRC at age 51 years with no family history of cancer who was identified by the universal screening of MMR proteins. Tumor tissue analysis showed loss of expression of MLH1 and PMS2, MSI, with absence of *MLH1* methylation and *BRAF* mutation. No germline mutation in *MLH1* was detected (Figure S3). Furthermore, analysis of somatic *MLH1* mutation in her tumor showed a homozygous/hemizygous nonsense pathogenic variant [c.1279C>T; p.(Glu427*)].

The fourth independent family with the same germline mutation in *POLD1* fulfilled Bethesda criteria. The proband was diagnosed with a CRC and a breast cancer at ages 45 and 52, respectively. Her father was found to have a CRC at age 83. Tumor analysis of the proband showed a normal expression of the MMR proteins and MSS (Figure S4).

In summary, 13 tumors diagnosed in 10 carrier patients (8 CRC, 3 endometrial cancers, a GIST and a breast cancer), an esophageal benign tumor and colorectal oligopolyposis (≤3 adenomas) in another 3 patients were considered. The median age of cancer onset for all *POLD1* mutation carriers was 48 years (range 23–80 years). The observed penetrance was 50% (40% in males and 53% in females). The cumulative risk at age 50 was 30% (40% in males and 27% in females) (Table 2).

DISCUSSION

To our knowledge, the 4 Valencian families included in this study are the only families where the *POLD1* Leu474Pro mutation has been found. The Valencian Community is located on the south-east Mediterranean coast of Spain, with a population of about 5 million people. The limited geographical dispersion of the 4 families, an area with an expanse of about 350 km, and the existence of a clear shared haplotype in carrier individuals, strongly suggests a common ancestor for these families. This mutation is unambiguously responsible for the high risk of cancer seen in these families. When we estimated the mutation age we obtained a wide range in the number of generations. This is because of the methodological limitation of the single marker method, especially when the distances of the first recombinant marker from both sides of the mutation are asymmetric. Therefore, these results must be viewed with caution. The *POLD1* Leu474Pro is the first founder mutation described in this gene to date. Likely, a broad incorporation of *POLE/POLD1* testing in the hereditary cancer diagnosis could unveil new founder mutations in these genes.

Founder mutations have been proven to impact upon molecular diagnosis strategies in specific populations; where they can be assessed as the first screening step, and if positive, avoid the expense of further gene scanning. This is valid even at a time when next generation sequencing (NGS) is been introduced in the molecular diagnosis of hereditary cancer syndromes. The inclusion of new genes or amplicons containing the founder mutations in the NGS gene panels makes possible a most effective and efficient identification of pathogenic mutations.[10]

Families 1 and 2 were classified as familial CRC type X and family 4 fulfilled Bethesda criteria with a MMR proficient tumor. Consequently, no further genetic studies could be offered for these 3 families according to the classical diagnostic algorithm for LS. Family 3 does not fulfil the Bethesda criteria and was referred to the Genetic Counseling in Cancer Unit for prospective universal screening of all CRC by IHC of MMR proteins. This case was classified as “Lynch-like syndrome” until further evidence could determine the somatic or hereditary origin of this tumor. The finding of the germline *POLD1* mutation made us wonder whether the loss of MLH1 expression could be the result of *MLH1* somatic mutations as a consequence of the

POLD1 inactivation. Effectively, we found a *MLH1* somatic homozygous/hemizygous nonsense mutation that could explain the loss of expression. Recently, Jansen et al., have published similar findings in another case and suggest that faulty proofreading may result in loss of MMR and thereby in MSI.[5] These investigators also emphasize the importance of *POLE/POLD1* germline and somatic screening in unexplained MMR-deficient tumors. The fact that germline mutations in non-MMR genes can mimic LS has been previously reported in *MUTYH.* Biallelic mutation in this gene is able to produce somatic mutations in MMR genes imitating the molecular and pathological phenotype of Lynch tumors.[11-12]

We acknowledge a limitation regarding to the clinical phenotype associated with this founder mutation. There is a bias in the *POLD1* studied population, because only unexplained cases of LS have been tested in our region. It is possible that some patients from the Valencian region with unexplained cases of MAP syndrome could also carry this founder mutation. The phenotype overlapping PPAP with LS and MAP syndromes, sustains broadly similar screening recommendations and management algorithms,[1,6] which are also suitable for this founder mutation.

New high-throughput technology is unveiling an increasing complexity of the underlying genetic background responsible for familial cancer syndromes. The discovery of a high level of genetic heterogeneity and variable expressivity, with a substantial phenotypic overlap among the defined syndromes, makes necessary a constant re-evaluation of the diagnostic approaches and algorithms, as well as the definition of new hereditary scenarios for cancer susceptibility. The PPAP syndrome is a good example of this, although in many cases polyposis is not the relevant phenotype of the syndrome.[1]

Leu474Pro is the first founder mutation in *POLD1* described to date. Future studies using NGS with large gene panels for any hereditary cancer condition will offer the possibility of detecting *POLE/POLD1* mutations in unsuspected clinical situations showing a more real and unbiased picture of the associated phenotype.

ACKNOWLEDGEMENTS

We are indebted to the patients and their families. We thank all members of the Hereditary Cancer Program of the Valencian Community (Spain).

This work was supported by the Institute for Health and Biomedical Research of Alicante (ISABIAL, UGP-16-146). RFA is recipient of a Fellowship from the Consellería Educación of the Valencian Community. ACA is funded by the Acción Juvenil from the Spanish Ministry of Economy and Competitiveness. VDO is recipient of a Fellowship from the Spanish Assosiation Against Cancer (AECC). AC and MIC are funded by Health and Biomedical Research Foundation from the Valencian Region (FISABIO). EHI is recipient of a fellowship from the Fondo Investigación Sanitaria ISCIII (FI12/00233).

CONFLICTS OF INTEREST

The authors of this article certify that they have no affiliations with or involvement in any organization or entity with any financial interest (such as honoraria; educational grants; participation in speakers' bureaus; membership, employment, consultancies, stock ownership, or other equity interest; and expert testimony or patent-licensing arrangements), or non-financial interest (such as personal or professional relationships, affiliations, knowledge or beliefs) in the subject matter or materials discussed in this manuscript.

AUTHORS' CONTRIBUTIONS

Conception and design: José-Luis Soto, Adela Castillejo.

Development of methodology: José-Luis Soto, Rosario Ferrer-Avargues, Virginia Díez-Obrero, Ester Martín-Tomás, Adela Castillejo, Eva Hernández-Illán, María–Isabel Castillejo, Alan Codoñer-Alejos, Víctor-M Barberá.

Acquisition of data: Ester Martín-Tomás, Ana-Beatriz Sánchez-Heras, Ángel Segura, María-José Juan, Isabel Tena.

Analysis and interpretation of data: Rosario Ferrer-Avargues, Virginia Díez-Obrero, Ester Martín-Tomás, Adela Castillejo, José-Luis Soto.

Writing, review, and/or revision of the manuscript: All authors.

Study supervision: José-Luis Soto, Adela Castillejo.

REFERENCES

1. Rayner E, van Gool IC, Palles C, et al. A panoply of errors: polymerase proofreading domain mutations in cancer. Nature Reviews Cancer. 2016;16(2):71-81. doi: 10.1038/nrc.2015.12.
2. Johnson RE, Klassen R, Prakash L, Prakash S. A Major Role of DNA Polymerase δ in Replication of Both the Leading and Lagging DNA Strands. Molecular Cell. 2015;59(2):163-75. doi: 10.1016/j.molcel.2015.05.038.
3. Cancer Genome Atlas Network. Comprehensive molecular characterization of human colon and rectal cancer. Nature. 2012;487(7407):330-7. doi: 10.1038/nature11252.
4. Palles C, Cazier JB, Howarth KM, et al. Germline mutations affecting the proofreading domains of POLE and POLD1 predispose to colorectal adenomas and carcinomas. Nature Genetics. 2013;45(2):136-44. doi: 10.1038/ng.2503.
5. Jansen AM, van Wezel T, van den Akker BE, et al. Combined mismatch repair and POLE/POLD1 defects explain unresolved suspected Lynch syndrome cancers. European Journal of Human Genetics. 2016;24(7):1089-92. doi: 10.1038/ejhg.2015.252.
6. Bellido F, Pineda M, Aiza G, et al. POLE and POLD1 mutations in 529 kindred with familial colorectal cancer and/or polyposis: review of reported cases and recommendations for genetic testing and surveillance. Genetics in Medicine. 2016;18(4):325-32. doi: 10.1038/gim.2015.75.
7. Valle L, Hernández-Illán E, Bellido F, et al. New insights into POLE and POLD1 germline mutations in familial colorectal cancer and polyposis. Human Molecular Genetics. 2014;23(13):3506-12. doi: 10.1093/hmg/ddu058.
8. Pérez-Carbonell L, Ruiz-Ponte C, Guarinos C, et al. Comparison between universal molecular screening for Lynch syndrome and revised Bethesda guidelines in a large population-based cohort of patients with colorectal cancer. Gut. 2012;61(6):865-72. doi: 10.1136/gutjnl-2011-300041.
9. Machado PM, Brandão RD, Cavaco BM, et al. Screening for a BRCA2 rearrangement in high-risk breast/ovarian cancer families: evidence for a founder

effect and analysis of the associated phenotypes. Journal of Clinical Oncology. 2007;25(15):2027-34.

10. Ponti G, Castellsagué E, Ruini C, Percesepe A, Tomasi A. Mismatch repair genes founder mutations and cancer susceptibility in Lynch syndrome. Clinical Genetics. 2015;87(6):507-16. doi: 10.1111/cge.12529.
11. Morak M, Heidenreich B, Keller G, et al. Biallelic MUTYH mutations can mimic Lynch syndrome. European Journal of Human Genetics. 2014;22(11):1334-7. doi: 10.1038/ejhg.2014.15.
12. Castillejo A, Vargas G, Castillejo MI, et al. Prevalence of germline MUTYH mutations among Lynch-like syndrome patients. European Journal of Cancer. 2014;50(13):2241-50. doi: 10.1016/j.ejca.2014.05.022.

Table 1. Minimum common haplotype.

The common haplotype is depicted, showing the minimal common region among the 4 families. All analyzed carriers of the c.1421T>C variant shared a common haplotype comprising the region between *D19S904* and *D19S921*. Minimum common haplotype: 2,995 Mb. Estimated frequency of the common haplotype in controls: 1 in 1.2 million.

Marker	D19S420	D19S902	D19S879	rs7246646	D19S867	D19S904	POLD1 c.1421T>C	D19S907	rs1363756	rs12975011	D19S206	D19S921	D19S927	D19S891	D19S887
Position (Mb)	43.305	47.829	49.013	49.140	50.035	50.273	50.406	50.558	50.632	51.348	52.048	53.268	53.795	55.522	57.122
Family 1	263	199	240	G	98	207	Yes	216	G	G	138	233	138	170	253/249
Family 2	263	213	244	G	104	207	Yes	216	G	G	138	233	138	184	255
Family 3	271	209	240	G	104	207	Yes	216	G	G	138	233	128	170	247
Family 4	271/257	209	240	G	104	207	Yes	216	G	G	138	233	128	170	249/255

2.995 Mb

3.233 Mb

	D19S867	D19S904	POLD1 c.1421T>C	D19S907	rs1363756	rs12975011	D19S206	D19S921	D19S927
Shared alleles	104	207	Yes	216	G	G	138	233	138
Control frequencies (%)	21.6	25	0	1.7	47	47	3.3	11.7	34.5

Table 2. **Clinical and histopathological information of *POLD1:* c.1421T>C carriers and obligated carriers.**

	IDENTIFICATION	SEX	AGE DIAGNOSIS	TUMOR TYPE AND POLYPS	LOCALIZATION	STAGE
FAMILY 1	F1-II-2	F	52	Adenocarcinoma	Endometrium	ND
	F1-II-6	F	39	Adenocarcinoma	Colon	IIA
			56	Adenocarcinoma	Endometrium	II
	F1-III-1	F	35	Adenocarcinoma	Sigmoid Colon	I
			35	GIST	Small Intestine	IA
FAMILY 2	F2-II-1	M	42	Adenocarcinoma	Colon	IV
	F2-III-2	F	48	Polyp (n=1)	Colon	-
			58	Adenocarcinoma	Endometrium	IB
	F2-III-4	M	50	Adenocarcinoma	Rectum	I
	F2-III-6	M	45	Polyp (n=1)	Colon	-
			47	Polyps (n=2)	Colon	-
	F2-IV-1	F	23	Adenocarcinoma	Left colon	I
			31	Benign tumor	Oesophagus	-
			31	Polyps (multiple)	Stomach	-
	F2-IV-2	F	25	Polyp (n=1)	Colon	-
			27	Polyp (n=1)	Colon	-
FAMILY 3	F3-II-6	F	80	Adenocarcinoma	Colon	IV
	F3-III-2	F	51	Adenocarcinoma	Cecum	IVB
FAMILY 4	F4-II-2	F	45	Adenocarcinoma	Rectum	III
			52	Adenocarcinoma	Breast	IIIA

ND, not determined.